便携式气象色谱仪与深度学习相结合的车内气味实时监测与评价

田丁炜 [1]

[1]（北京化工大学化学学院 北京 100029）

摘要

[目的] 本研究的目标是开发一种便携式气相色谱仪，结合机器学习实现现场的 VOC 采集和快速的气味评价。

[方法] 我们使用卷积神经网络-长短期记忆（CNN-LSTM）建立了气味强度的预测模型。由于收集的数据量较小，我们使用生成对抗网络（GAN）对每个气味强度类别的 VOC 数据进行了生成，以增强模型的训练。

[结果] 在生成数据后，我们再次使用 CNN-LSTM 建立了模型，并与人工神经网络（ANN）、支持向量机（SVM）和梯度提升决策树（XG-Boost）进行了比较。结果表明，使用 GAN 生成数据后的测试准确率优于原始数据。

[局限] 未来的工作将集中在进一步优化模型和扩大数据集上，以提高预测的准确性和稳定性。

[结论] 这项研究表明，通过使用深度学习和生成对抗网络，我们可以有效地预测车内的气味强度，从而改善车内的空气质量。此外，我们还将探索将这种方法应用于其他环境条件下的空气质量预测。这为未来的空气质量监测和改善提供了新的可能性。由于我们的设备是便携式的，且模型结构较小可以直接嵌入到设备上，从而实现现场的 VOC 采集和快速的气味评价。这为未来的空气质量监测和改善提供了新的可能性。

关键词：便携式气相色谱，现场 VOC 采集，气味强度预测，深度学习

分类号：TP393

Real-time Monitoring and Evaluation of Indoor Odor with Portable Gas Chromatograph Combined with Deep Learning

Dingwei Tian [1]

[1](Affiliation: School of Chemistry, Beijing University of Chemical Technology, Beijing 100029)

Abstract:

[Objective] The aim of this study is to develop a portable gas chromatograph, combined with machine learning, to achieve on-site VOC collection and rapid odor evaluation.

[Methods] We used a Convolutional Neural Network-Long Short-Term Memory (CNN-LSTM) to establish an odor intensity prediction model. Due to the small amount of data collected, we used a Generative Adversarial Network (GAN) to generate VOC data for each odor intensity category to enhance model training.

[Results] After generating the data, we used CNN-LSTM to establish the model again and compared it with Artificial Neural Networks (ANN), Support Vector Machines (SVM), and Gradient Boosting Decision Trees (XG-Boost). The results showed that the test accuracy after using GAN to generate data was better than the original data.

[Limitations] Future work will focus on further optimizing the model and expanding the dataset to improve the accuracy and stability of the prediction.

[Conclusion] This study shows that by using deep learning and generative adversarial networks, we can effectively predict the odor intensity inside the car, thereby improving the air quality inside the car. In addition, we will explore the application of this method to air quality prediction under other environmental conditions. This provides new possibilities for future air quality monitoring and improvement. As our equipment is portable and the model structure is small enough to be directly embedded into the device, it can achieve on-site VOC collection and rapid odor evaluation. This provides new possibilities for future air quality monitoring and improvement.

Keywords: Portable Gas Chromatography, On-site VOC Collection, Odor Intensity Prediction, Deep Learning

1. 背景

车内的挥发性有机物（VOC）是影响车内空气质量的主要原因之一。主要包括烷烃、烯烃、芳烃和醛酮类物质[1]。这些 VOC 由车内饰件释放出来后，会对人体产生不同程度的危害。在日晒导致的高温下，车内零部件及装饰材料中的 VOC 则更易挥发出来[2,3]。根据 J.D. Power 在中国开展的 "初始质量研究"（IQS），"

车内异味 "问题一直是新车车主的一个突出问题。在 2023 年的研究报告中，该问题被列为内燃机汽车和新能源汽车车主遇到的 20 大问题之一。 值得注意的是，在过去三年中，这一问题一直排在前三位[4]。因此，为了满足用户需求，提高各主机厂的产品竞争力，从实际出发，对车内异味、零部件异味、材料异味进行评估至关重要。

目前，汽车主机厂普遍采用主观气味评估方法来进行气味的质量控制，该方法直接反映了用户对气味的接受程度[5,6]。该方法是由专业的气味评价员组成的团队来进行主观气味评级。然而，这种方法存在一些限制。首先，评估结果容易受到评估人员个人身体和心理状态的影响。此外，不同个体之间的评估结果存在差异，如性别差异等[7]。其次，不同气味评估组之间的评估结果可能不一致，同一材料在不同评估机构的评估中可能得出不同的结果[8]。同时，高温处理的材料可能释放大量有害的挥发性有机化合物，持续的气味评估工作可能对评估人员的长期健康构成风险。因此，建立客观的气味评估方法具有至关重要的意义。

当前，开发客观的气味评估体系已经成为了环境研究领域的一个焦点[9]。现阶段，主要的方法是基于 Weber-Fechner law 将某几种 VOC 与气味的感官强度级别联系起来。具体来说，该方法是在一个已知的体系中加入某几种固定的 VOC，将其测定的浓度与这个体系的气味进行关联[10]。

然而，气味相互作用的情况却普遍存在，且会受到多种作用的影响，多种 VOCs 之间会产生相互作用导致其混合物的气味强度会低于简单的各种 VOC 单独体系气味强度的叠加[11]。这对于车内复杂的 VOC 体系气味的客观评估则造成了困难。然而使用机器学习（ML）则可以自发的学习 VOCs 与气味之间的关系，因此可以用于气味强度建模研究。

为了实现现场快速气味客观评价，将 VOCs 进行快速的采集和浓度计算则是有必要的。便携式的气相色谱仪（PGC）由于其体积小，重量轻，方便携带。这使得它可以在任何地方进行现场检测，无需将样品带回实验室。同时气相色谱可以在短时间内对各种待检测的 VOC 实现分离并且使用检测器各自进行定量。从而能够实现现场快速的 VOC 定量分析[14]。光离子化检测器(PID)的工作原理是使用一个紫外光源(通常是 10.6 eV 的氙灯)去电离检测气体样本分子。当样本分子吸收紫外光的能量时,如果该能量高于分子的电离能,样本分子会被电离形成正负离

子。这些电离后的离子会导致测量电流的变化,从而检测和定量样品中组分的浓度[12]。PID 具有体积小、灵敏度高、响应快等优点。结合上述优势，可以使用 GC-PID 来实现 VOC 的现场快速检测。

本研究的目的是基于深度学习方法开发一种，现场快速气味评价的方法。具体来说，我们使用 GC-PID 采集了现场车内的 VOCs 数据和气味强度数据 116 条。LSTM 具有记忆功能可以把握各种 VOC 之间的相互联系和作用[13]。因此，我们使用采集 VOC 和气味的数据基于 CNN-LSTM 建立了气味预测模型。由于各个气味类别的数据分布并不均匀，因此我们使用生成对抗网络（GAN）对每个气味强度类别的 VOCs 数据进行了生成，以增强模型的训练。为了选择最佳的 ML 算法来建立当前的气味模型，我们使用了人工神经网络（ANN）、支持向量机（SVM）和梯度提升决策树（XG-Boost）与 CNN-LSTM 进行比较。

2. 材料和方法

2.1 VOCs 和气味数据采集

我们的数据集包括汽车内饰件和整车一共 116 条数据。对于内饰件来说，我们在采集 VOCs 时主要参考法规 ISO 12219-9[15]中规定的采样袋方法。首先将内饰件同表面的 PE(聚乙烯)软包装仪器在常温平衡仓中（T:25℃，RH:50%）平衡 24 h。将内饰件放入 2000 L Tedlar 采样袋(Ningbo Huance Equipment co., LTD.)对材料进行测试。在确保采样袋充分老化后，将样品于采样袋中密封，使用 30% N_2 清洗采样袋三次后充入 50% N_2，使用高温测试仓将采样袋在 65±2℃下平衡 2 h，使得 VOCs 释放平衡。随后，温度保持在 65℃时进行采样。

在整车的研究中，我们参照了中国环保行业标准 HJ/T 400-2007[16]的规定，对整车进行了 VOCs 采集。然而，由于我们进行的是现场快速测试，采集过程并未严格按照标准中规定的在 25℃的条件下封闭 16 小时。相反，我们根据实际的工况，对车内环境进行了平衡和采集。这种方法虽然与标准有所不同，但更贴近实际的使用环境。

我们使用便携的 GC-PID 进行 VOCs 采集和定量，采集流速为 40 mL/min，采集时间 2 min。色谱柱：AB-INOWAX(30 m ×0.32 mm×1 μm)，载气为 N_2, 程序升

温如**表 1** 所示：

表 1　GC-PID 程序升温条件

梯度	速率 ℃/min	温度℃	保持时间 min
1		40	2.2
2	2.58	55	0
3	4.17	60	0
4	2	65	4
5	38	160	0.3
6	5	220	10

气味强度的评估是由专业的气味评价团队进行的，该团队由 5 到 6 名经过专业训练的气味评价员组成。我们采用了一个六级的气味强度评价系统，等级从 1（无法检测）到 6（无法忍受）。原则上，评价员们以 0.5 的等级为间隔进行评价，但为了获得更精确的评价结果，我们也允许评价员在评价时给出 3.0+或 3.0-的结果。这种评价方法不仅有利于我们的模型进行更精确的分级，而且可以有效地监督气味是否超过标准，同时还为模型提供了更大的容错空间。所有评价员给出的气味强度的算术平均值应作为最终的气味结果 [17]。对于内饰件来说，在 VOCs 采样完成后，将 2000 L 采样袋边缘的密封条打开一个约 4-6 cm 的小口，由 6 名气味评价员依次评价。对于整车来说，在 VOCs 采集完成后，5 名气味评价员打开车门并在 10 s 之内上车进行气味评价。

2.2 ANN

人工神经网络（ANN）是模拟人脑行为的计算模型，用于学习输入和输出之间的关系。我们使用反向传播（BP）算法训练 ANN，通过计算输出误差并反向传播来调整参数，使输出接近期望值。在本研究中，我们使用 BP 训练 ANN 进行气味强度分级。ANN 的结构如图 1 所示：

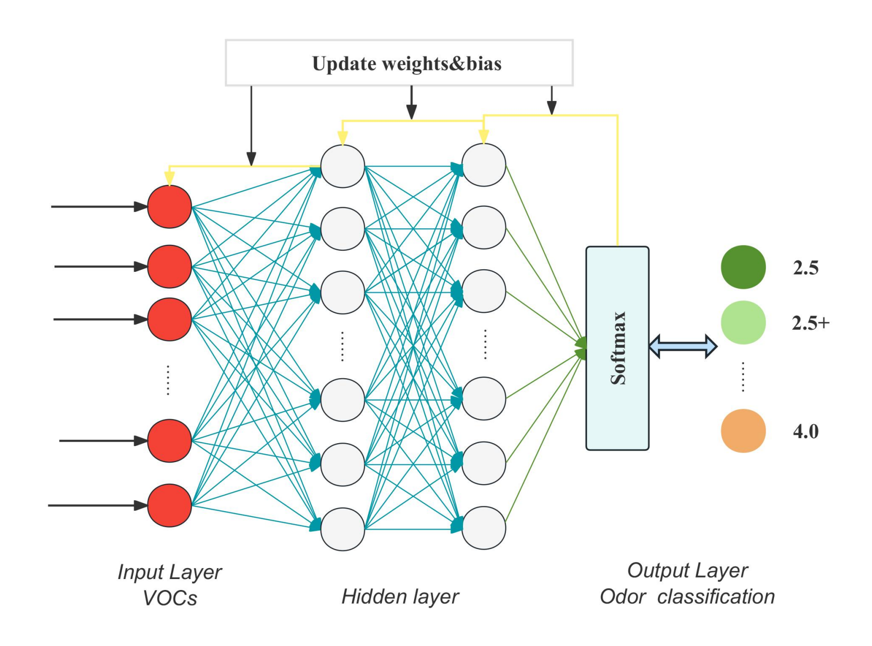


图 1 ANN 气味模型结构

如图所示，我们使用了一个双隐藏层结构的 ANN 网络，VOCs 数据经过隐含层计算后回归到 C 个维度（C 代表气味强度类别,C=10）并通过 Softmax 函数处理将多分类的输出值转换为范围在[0, 1]和为 1 的概率分布，从而得到各个气味强度级别的概率。所使用的损失函数为 CrossEntropyLoss，我们使用了 AdamW 作为优化器并加入权重衰减系数（Weight_Decay），它通过在损失函数中添加正则化项，惩罚模型的权重参数，使模型更倾向于选择较小的权[18]. 通过控制 Weight_Decay 值的大小，可以有效的对模型的过拟合和欠拟合问题进行权衡，优化出最佳的模型。设置学习率为 0.01，Weight_Deca 设为 0.1。将数据集划分为训练集，验证集和测试集三部分。由于各种 VOC 之间的绝对值差异很大，VOCs 数据在输入网络前进行了 Z-Score 处理，它会将每种 VOCs 进行标准化，使得数据均值为 0，方差为 1。训练时采用了一种基于验证集的早停策略。具体来说，在每次训练迭代过程中都会检查模型在验证集上的准确率。一旦模型在验证集上的准确率达到目标就会停止训练并保存当前的模型参数并用于测试。

2.3 XG-Boost 和 SVM

XG-Boost 和 SVM 已经被广泛的应用于各种分类问题中[19,20]。对于 SVM 来说，由于其本质是对于二元分类问题选择最优的分类超平面，因此其无法直接处理多分类问题。通常 SVM 会以"一对一"或"一对多"策略来进行多分类[21]。在本研究中 SVM 采取"一对一"策略来进行多分类。同时，为了寻找到最佳的超参数组合，

我们使用了网格搜索（GridSearch）算法来对模型的超参数寻优。

2.4 CNN-LSTM

本研究构建了基于 CNN 进行特征增强，LSTM 进行记忆从而把握把握各种 VOC 之间相互联系的气味分类模型。具体来说数据以[1×n]的维度输入到三层卷积层中，其中 n 是指 VOC 的种类数。对每一层使用 16k 个卷积核进行步长为 1 的卷积（k 为层数=1，2，3）。这样，经过三层卷积后数据的维度就会变成（1×64×n）。在每一层卷积运算完成后，我们都接上了一个 BatchNorm 层，并且使用 Relu 激活函数对数据进行非线性变换。随后数据经过一个神经元个数为 64 的全连接层，这一层不会改变数据的维度却有助于提高模型的表达能力。随后在经过一个 Dropout 层处理后数据以(n×64)的形式输入到 BiLSTM（双向 LSTM）层中，n 为 seq_len，64 为 input_size，这代表每种 VOCs 以 64 维的向量进行表示。这表示在每个时间步（1-n），BiLSTM 会读取每个序列的一种 VOCs（转化为 64 维的向量），并更新其隐藏状态。这样 LSTM 则会在前后两个方向捕捉 VOCs 浓度之间的相互影响。随后经过一个全连接层并将数据维度回归到（1×10）经过 Softmax 层转化为各个浓度级别的概率并输出。算法的优化器和损失函数等配置同 AdamW 优化器。

2.5 GAN

在我们的研究中，我们采用了 CNN-LSTM 深度学习模型。尽管该模型是轻量级的，但仍需要大量训练样本才能充分发挥其性能。为解决这一问题，我们采取了增加训练集样本量的策略，并尽可能使各类别标签分布均匀。然而，实际训练样本中，气味强度≥3.7 的零部件或整车样本较少，这些是我们需要重点关注的对象。因此，我们采用了生成对抗网络（GAN）基于训练集数据对各气味类别的 VOC 进行数据生成。

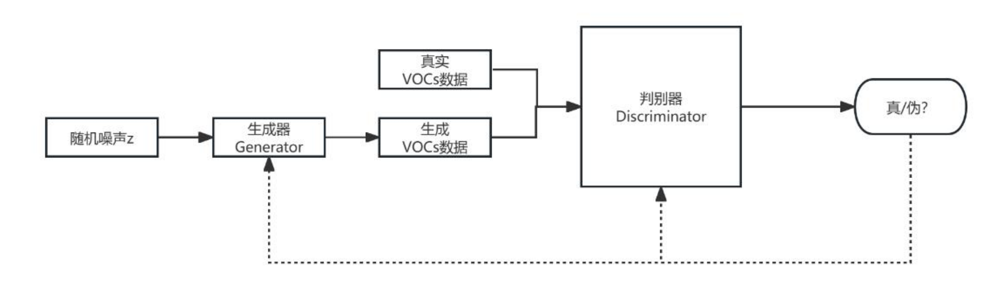


图 2　GAN 网络的原理图

在数据进行预处理之后，GAN 首先定义了判别器和生成器模型。接着，通过对抗训练的方式，训练判别器和生成器，使其生成逼真的数据样本。最后，使用生成器生成新的数据，并与原始数据拼接在一起。这种数据增强方法可以对每个气味强度下的 VOCs 数据进行生成，扩充数据集的规模和多样性从而提升模型的泛化能力和性能。GAN 的原理如图 2 所示，使用 GAN 进行数据生成并建立模型的流程如图 3 所示。

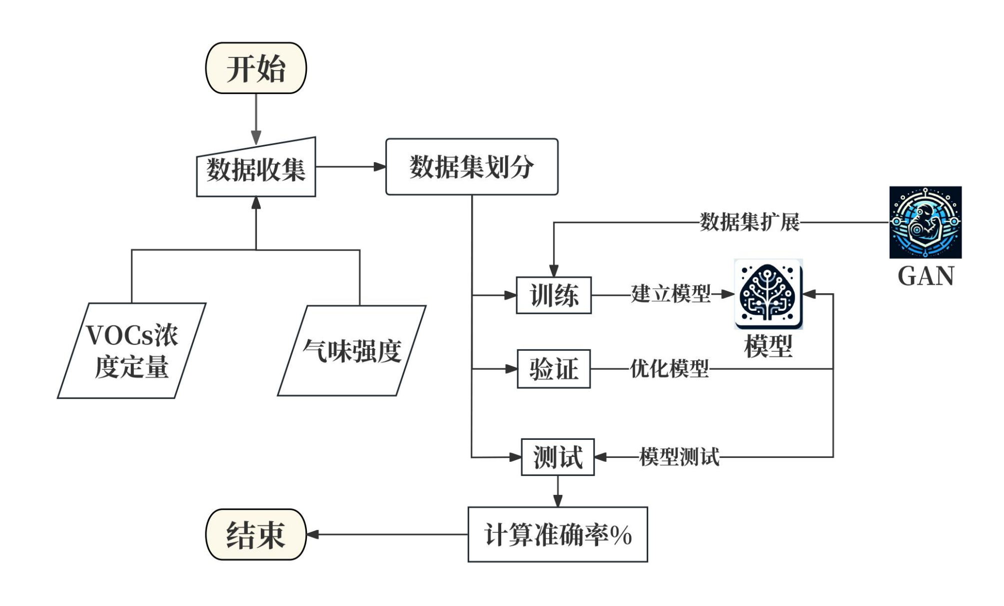


图 3 使用 GAN 进行数据生成并建立模型的流程图

3. 结果与讨论

3.1 数据描述

在本研究中，共收集了 118 条零部件和整车 VOCs 数据以及对应的气味标签，涵盖了 10 个气味类别。如表 2 所示，各个类别的气味标签分布并不均匀，特别是 3.5 级别以上的气味样本数量较少。这主要是因为在实际测试中，气味强度大于 3.5 的样本很少见，尤其是 4.0 级别以上的气味样本几乎不存在。因此，为了解决这个问题，采用了生成对抗网络（GAN）进行数据增强，将生成的数据与原始数据一起用于模型训练。这样做的目的是为了解决数据量分布不均匀可能导致的过拟合和欠拟合问题，从而提高模型的泛化能力。在使用 GAN 生成数据后，各个气味类别的数据分布在表 2 中进行了总结，生成后的数据分布更为均匀。然而，由于个别级别的样本量较少，GAN 的学习对象有限，如果生成大量数据，可能会导致生成的数据过分依赖于个别数据，从而造成严重的过拟合。

表 2 各个类别气味标签的分布情况

气味类别	数据（原始）	数据（增强后）
2.5	9	32
2.5+	3	18
3.0-	19	34
3.0	33	35
3.0+	29	29
3.5-	4	32
3.5	16	33
3.5+	3	21
4.0-	1	11
4.0	1	11

我们使用 PGC 收集的 VOCs 共有 8 种（其中包括 5 种苯系物和 3 种醛类化合物），并且我们收集了 TVOC（总 VOCs），其中 TVOC 是甲苯进行半定量的，因此模型的输入维度为 [1×9]，模型的输出则为气味强度类别。

3.2 未使用 GAN 时的模型

当我们使用原始的 118 条数据建模时，我们使用了 93 条数据进行建模，各自使用 14 条数据来进行验证和测试。由于 4.0-和 4.0 的数据只有一条，因此在验证和测试的数据中则没有包含这两个类别。表 3 总结了各个算法的准确率和参数。除了计算模型对气味强度的完全匹配准确率外，我们还计算了在气味强度修约到 0.5 级后的准确率，这是一个衡量模型鲁棒性的重要指标。在实际的气味评价过程中，气味强度通常会被修约至 0.5 级，因此，这种修约后的准确率能够更好地反映模型在实际应用中的性能。图 4 则更佳准确的展示出各种算法在验证和测试集各个气味类别上的预测能力。

表 3 原始训练集各模型的准确率和参数

Algorithm \	Accuracy					Parameters
Data set	Train set	Val set	Val/0.5	Test set	Test/0.5	
ANN	96.67%	50.00%	92.86%	35.72%	92.86%	Lr:0.01, W_D:0.1, hidden nodes:(11,22), epoch:110

SVM	51.61%	28.57%	92.86%	21.43%	92.86%	Kernel:rbf, C:0.1, gamma:1
XG-Boost	72.04%	57.14%	92.86%	7.14%	78.57%	Lr:0.01, max_depth:5, n_estimators:200
CNN-BiLSTM	88.17%	57.14%	100%	42.86%	100%	Conv:16-32-64, lr:0.001, W_D:0.1, epoch:60 BiLSTM:150, Fc:(64,150), Dropout:0.2

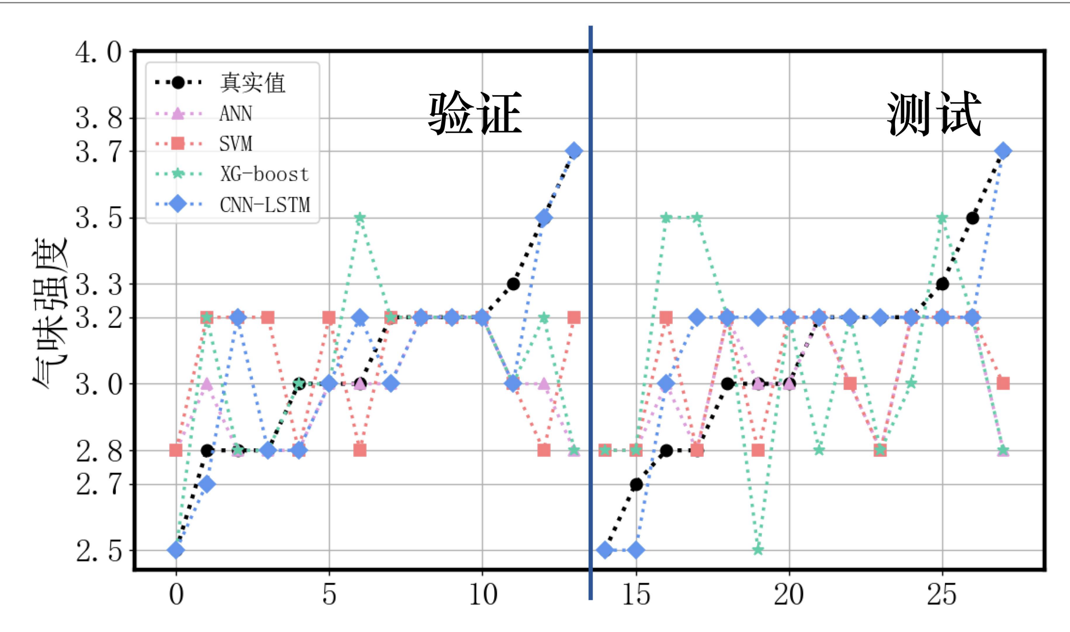


图 4 未使用 GAN 时，各个模型的验证和测试结果

从表 4 可以看出，ANN 和 CNN-LSTM 在训练集上都取得了较好的结果，在验证集和测试集上 CNN-LSTM 略好过 ANN，且 CNN-LSTM 在 0.5 级以内的准确率为 100%，这表面 CNN-LSTM 模型已经能在 0.5 级内准确的预测气味强度。

3.2 使用 GAN 后的模型

在使用了 GAN 进行训练数据增强后，训练数据共 228 条，为了能够横向比较数据增强前后模型的结果，验证和测试数据保持不变。

表 4 训练集使用 GAN 后各模型的准确率和参数

Algorithm \	Accuracy					Parameters
Data set	Train set	Val set	Val/0.5	Test set	Test/0.5	
ANN	94.30%	50%	50%	100%	100%	Lr:0.01, W_D:0.1, hidden nodes:(11,22), epoch:110

SVM	96.05	35.71%	35.71%	85.71%	85.71%	Kernel:rbf, C:0.1, gamma:1
XG-Boost	100%	42.86%	100%	21.43%	78.57%	Lr:0.01, max_depth:5, n_estimators:200
CNN-BiLSTM	95.18%	78.57%	100%	57.14%	100%	Conv:16-32-64, lr:0.001, W_D:0.1, epoch:60 BiLSTM:150, Fc:(64,150), Dropout:0.2

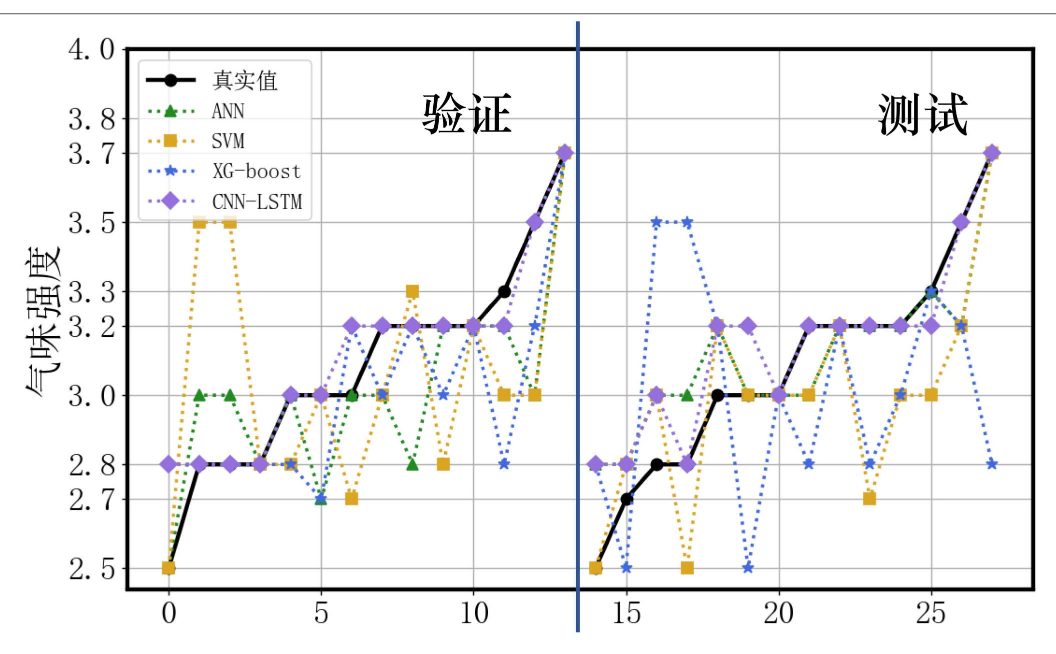


图 5 GAN 增强训练集后各模型的验证和测试结果

由表 4 可知，GAN 对训练集数据增强后，各个模型从验证和测试准确率皆有提升，CNN-LSTM 的验证准确率达到了 78.57%（11/14）。测试集准确率也达到了 57.14（8/14）%。0.5 级之内的准确率均为 100%，此时模型已经能够准确的对气味进行预测，其泛化能力还需要进一步提升。值得注意的是此时的 CNN-LSTM 模型能对气味强度大于 3.5 的高气味强度数据吉奥为准确的判别，这对于气味识别监测异常件的作用则较为重要。

为了验证模型学习到了真实的规律，我们还使用了 SHAP（SHapley Additive Explanations）方法来对 CNN-LSTM 模型进行解释，以验证模型是否学习到了真正的气味规律。气味阈值是指人类能够察觉到某种物质气味的最低浓度。因此，气味阈值越低，意味着这种物质的气味在较低的浓度下就能被人类察觉到。换句话说，气味阈值越低的物质，其气味通常被认为更强烈。我们可以看出，各种 VOC 对气味的贡献并不是完美的符合嗅阈值的规律 [22]，嗅阈值的排序如表 5 所示。这可能是由于三方面的原因导致的，首先由于便携式设备对 VOC 的分离和监测能力有限，定量结果可能存在误差，且收集到的 VOC 种类较少。

其次单一 VOC 最终对气味产生影响的气味活度值还于 VOC 的浓度有关，不只依赖嗅阈值。最后收集到的样本量依然较小，不足以反应真实的气味规则。

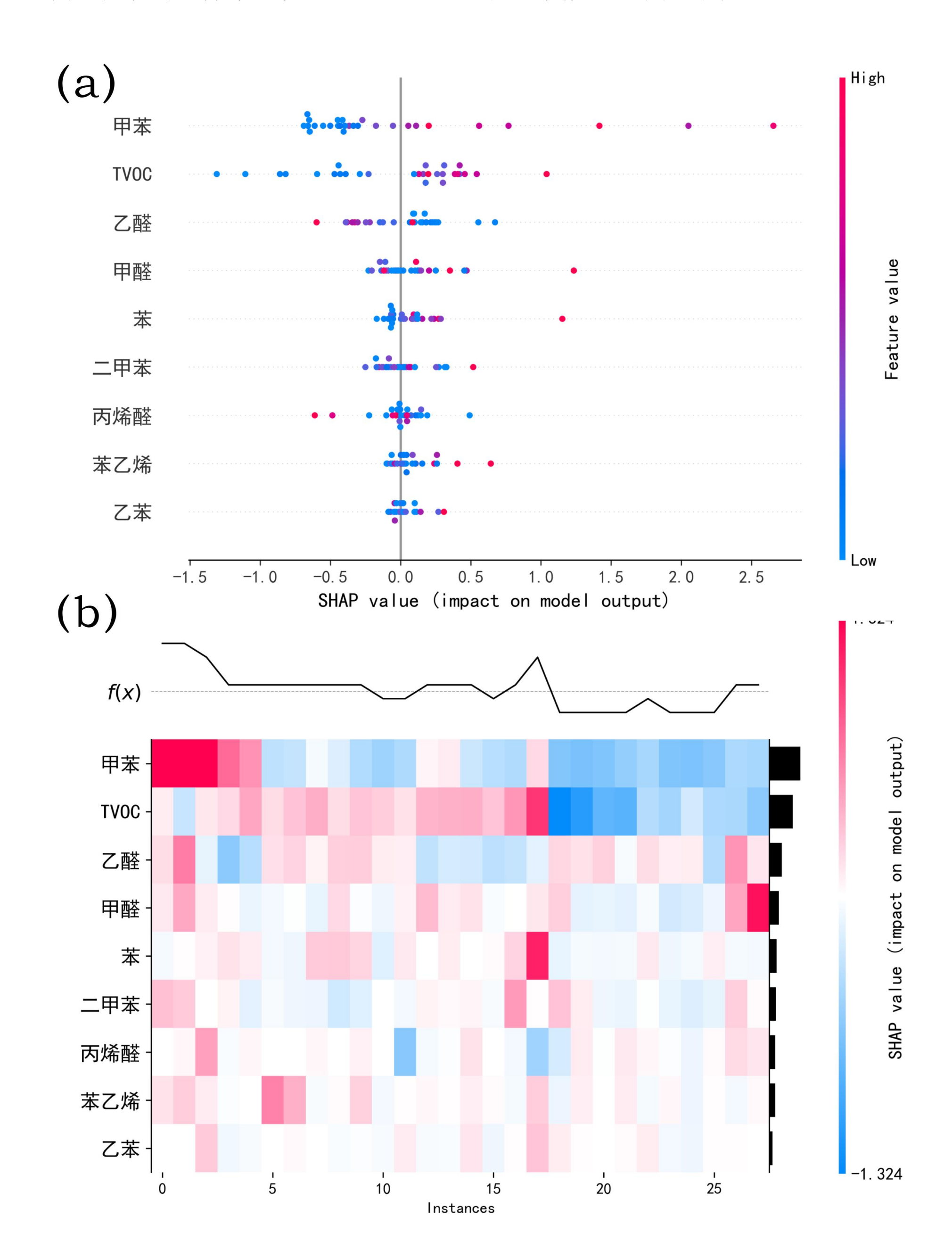


图 6 GAN 增强训练集后各模型的验证和测试结果

VOC **（按照嗅阈值排序）**	**嗅 阈 值** /ppm (10^-6 v/v)	**VOC 按照 SHAP 排序**
\	\	
乙醛	0.00154	甲苯
丙烯醛	0.0036	乙醛

苯乙烯	0.0350	甲醛
二甲苯	0.1596	苯
乙苯	0.1700	二甲苯
甲苯	0.3300	丙烯醛
甲醛	0.5000	苯乙烯
苯	2.7000	乙苯

4. 结论与展望

本研究基于 CNN-LSTM 和 PGC 开发了一种现场快速气味评价方法。该方法能够有效地对车内气味强度进行预测，10 个类别的气味识别准确率能达到 67.86%，且能够实现半级之内准确率为 100%的气味强度评价。与其他三种机器学习算法相比，CNN-LSTM 模型具有更好的泛化能力和鲁棒性。本研究为现场快速气味评价提供了新的思路和方法。未来，可以从以下几个方面进一步改进和完善该方法。首先可以进一步扩大数据集，以提高模型的泛化能力。可以将该方法与其他气味评价方法相结合，以提高气味评价的准确性和可靠性。也可以将该方法应用于其他需要且识别的领域，如食品、化妆品和纺织品等行业的气味评价。

参考文献：

[1]Yoshida T, Matsunaga I. A case study on identification of airborne organic compounds and time courses of their concentrations in the cabin of a new car for private use[J]. Environment International, 2006, 32(1): 58-79.

[2] Chien Y C. Variations in amounts and potential sources of volatile organic chemicals in new cars[J]. Science of the total environment, 2007, 382(2-3): 228-239.

[3] Geiss O, Tirendi S, Barrero-Moreno J, et al. Investigation of volatile organic compounds and phthalates present in the cabin air of used private cars[J]. Environment international, 2009, 35(8): 1188-1195.

[4] J.D. Power. (2023). China Initial Quality Study (IQS). J.D. Power website. 取自

[5] International Organization for Standardization. (2017). Odor determination in interior air of road vehicles and test chamber air of trim components by olfactory measurements (ISO 12219-7:2017). 取自 https://www.iso.org/standard/62641.html

[6] German Association of the Automotive Industry. (2018). Determination of the odor characteristics of trim

materials in motor vehicles (VDA 270-2018). 取自 https://webshop.vda.de/VDA/en/vda-270-06-2018

[7] Shanahan L K, Kahnt T. On the state-dependent nature of odor perception[J]. Frontiers in Neuroscience, 2022, 16: 964742.

[8] Ryu H W, Han J S, Choi S H, et al. Effect of panelist's general properties on the evaluation of complex odors[J]. Korean Journal of Odor Research and Engineering, 2012, 11(4): 151-158.

[9] Zhou Y, Vitko T G, Suffet I H M. A new method for evaluating nuisance of odorants by chemical and sensory analyses and the assessing of masked odors by olfactometry[J]. Science of The Total Environment, 2023, 862: 160905.

[10] Yan. L, Liu. J, Jiang. S, Wu. C, Gao. K, The regular interaction pattern among odorants of the same type and its application in odor intensity assessment, Sensors 17(7) (2017) 1624. https://doi.org/10.3390/s17071624

[11] Cain W S, Drexler M. Scope and evaluation of odor counteraction and masking[J]. Annals of the New York Academy of Sciences, 1974, 237(1): 427-439.

[12] Zhang W, Li H, Zhang Y, et al. Fast determination of monocyclic aromatic hydrocarbons in ambient air using a portable gas chromatography–photoionization detector[J]. Chromatographia, 2017, 80: 1233-1247.

[13] Bengio Y, Simard P, Frasconi P. Learning long-term dependencies with gradient descent is difficult[J]. IEEE transactions on neural networks, 1994, 5(2): 157-166.

[14] Frausto-Vicencio I, Moreno A, Goldsmith H, et al. Characterizing the performance of a compact BTEX GC-PID for near-real time analysis and field deployment[J]. Sensors, 2021, 21(6): 2095.

[15] International Organization for Standardization. (2019). Interior air of road vehicles — Part 9: Determination of the emissions of volatile organic compounds from vehicle interior parts — Large bag method (ISO 12219-9). 取自 https://www.iso.org/standard/75242.html

[16] Environmental Protection Industry Standard of China. (2007). Determination of Volatile Organic Compounds and Carbonyl Compounds in Cabin of Vehicles (HJ/T 400-2007).

[17] International Organization for Standardization. (2017). Odor determination in interior air of road vehicles and test chamber air of trim components by olfactory measurements (ISO 12219-7:2017).

[18] Loshchilov, I., & Hutter, F. (2017). Decoupled weight decay regularization. arv preprint arv:1711.05101.

[19] Chen T, Guestrin C. Xgboost: A scalable tree boosting system[C]//Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining. 2016: 785-794.

[20] Saigal P, Khanna V. Multi-category news classification using support vector machine based classifiers[J]. SN Applied Sciences, 2020, 2(3): 458.

[21] Chmielnicki W, Stapor K. Combining one-versus-one and one-versus-all strategies to improve multiclass SVM classifier[C]//Proceedings of the 9th International Conference on Computer Recognition Systems CORES 2015. Springer International Publishing, 2016: 37-45.

[22] Nagata, Y., 2011. Compilations of odour threshold values in air, water and other media (second enlarged and revised edition). Tokyo: Japan Association on the Environment of Odor. [https://www.academia.edu/3443692/Compilations_of_odour_threshold_values_in_air_water_and_other_media_second_enlarged_and_revised_edition_]

(通讯作者: 田丁炜　2021200994@buct.edu.cn)

作者贡献声明[¯]：

田丁炜：数据收集、数据的分析、论文起草、检查和审核。